ISO/IEC JTC1/SC2/WG2 N3120
L2/06-257
2006-08-06

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposal for encoding the Cham script in the BMP of the UCS**
Source: **UC Berkeley Script Encoding Initiative (Universal Scripts Project)**
Author: **Michael Everson**
Status: **Individual Contribution**
Replaces: **N1559, N1578, N1599, N1960**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2006-08-06**

1. Introduction. Cham is a Austronesian language of the Malayo-Polynesian family. The Eastern Cham number about 73,000, and live in Bình Thuận, Ninh Thuận, and Đồng Nai provinces in southern Vietnam, as well as in Hồ Chí Minh City. The Western Cham number about 25,000 in An Giang and Tây Ninh provinces, in Hồ Chí Minh City; in Cambodia, however they are more numerous, numbering 220,000. The Cham script is used more by the Eastern Cham. In this document, transliteration in *italics* follows that of Bùi Khánh Thế's 1995 Cham-Vietnamese dictionary; character names are given in SMALL CAPS.

2. Structure. The Cham script is of the Brahmic type. Consonants bear an inherent vowel (*-a* in the case of most consonants, *-ư* in the case of nasal consonants), but no virama is used to kill this vowel; vowel matras modify it, and explicit final consonants are used where there is no inherent vowel. Consonant conjuncts are not formed productively. A syllable is structured (and represented in the backing store) as (I | (C M? M?)) P? V? L? F?, where the first character is either an independent vowel "I" or a consonant "C" (which may also include the independent vowel ꨀ A since it can bear medial *-ꞃa* or *-wa* as discussed below); C may be followed by one or two optional medials M, and (I | (C M? M?)) may be followed by an optional pre-vowel P that is rendered before the consonant cluster, an optional vowel V, an optional vowel lengthener L, and optionally a final consonant F.

3. Independent vowel letters. Six of the initial vowels are represented with the unique independent vowels ꨀ A, ꨁ I, ꨂ U, ꨃ E, ꨄ AI, ꨅ O; ; these always indicate a syllable initial vowel, but they may occur word-internally at a syllable break: ꨕꨚꨚꨆ *ta'uk*. Other vowels which do not have independent forms are represented by vowel matras applied to the letter ꨀ A; ꨀꨩ *ā*, ꨀꨪ *i*, ꨀꨫ *u*, ꨀꨬ *ơ*, ꨀꨮ *o*, ꨀꨯ *ai*, ꨀꨮꨬ *au*, ꨀꨰ *ư* are attested in initial position in Bùi Khánh Thế 1995. Note that there is some redundancy in representation: ꨀꨪ and ꨁ both represent *i*, ꨀꨫ and ꨂ *u*, ꨀꨯ and ꨄ *ai*, ꨀꨮ and ꨅ *o*; all of these are unique spellings. I suggest it is prudent that applications permit ꨀ A to bear *any* of the vowel signs. Four of the other independent vowels are also attested bearing matras: ꨁꨪ *i* (I with redundant -I); ꨂꨪ *ū* (U with -AA as a vowel lengthener); ꨄꨮ *o* (AI with -O as a vowel modifer in ꨚꨄꨮꨚ *ppa'on*), ꨄꨯ *ai* (AI with redundant -AI), ꨄꨮꨬ *e* (AI with -O and -OE as vowel modifers), ꨅꨩ *ō* (O with -AA as a vowel lengthener).

4. Medial consonant signs Consonants can be followed by the glides ◌ꨚ -YA, ◌ꨛ -RA, ◌ꨜ -LA, and ◌ꨫ-WA, the first of which normally ligates with the consonant it modifies. When it does so, the vowel matra is applied to it (and not to the base consonant): ꨙꨚꨪ *hꞃā*. The independent vowel ꨀ A can bear two of the medials: ꨀꨚ *ꞃa* [ja] and ꨀꨫ*wa* [wa] contrast with ꨚ *ya* [ja] and ꨛ *va* [wa]. Three medial

clusters occur: ◌ྲྭ *-rwa*, ◌ྱ *-l̤ia*, and ◌ླྭ *-lwa*. The following combinations occur in word-initial position:

◌ྱ -YA: ཨྱ AYA, ཀྱ KYA, ཁྱ KHYA, གྱ GYA, ངྱ NGYA, ཅྱ CHYA, ཆྱ CHHYA, ཇྱ JYA, ཉྱ NHYA, ཏྱ TYA, ཐྱ THYA, དྱ DYA, ཉྱ DHYA, ནྱ NYA, ཌྱ DDYA, པྱ PYA, ཕྱ BYA, བྱ BHYA, མྱ MYA, ཙྱ BBYA, ཨྱ RYA, ལྱ LYA, ཤྱ SSYA, སྱ SYA, ཧྱ HYA

◌ྲ -RA: ཀྲ KRA, གྲ GRA, ཅྲ CHRA, ཇྲ JRA, ཉྲ NHJRA, ཏྲ TRA, ཐྲ THRA, དྲ DRA, ནྲ NRA, པྲ PRA, ཕྲ BRA, མྲ MRA, ཙྲ BBRA, ཤྲ SSRA, སྲ SRA, ཧྲ HRA

◌ྲྭ -RWA: ཀྲྭ GRWA, ཏྲྭ TRWA, དྲྭ DRWA, པྲྭ PRWA,

◌ླ -LA: ཀླ KLA, ཁླ KHLA, གླ GLA, པླ PLA, བླ BLA, ཙླ BBLA, ཧླ HLA

◌ླྱ -LYA: ཀླྱ KLYA

◌ླྭ -LWA: ཀླྭ KLWA, པླྭ PLWA, བླྭ BLWA, ཙླྭ BBLWA

◌ྭ -WA: ཨྭ AWA, ཀྭ KWA, ཁྭ KHWA, གྭ GWA, གྷྭ GHWA, ངྭ NGWA, ཅྭ CHWA, ཆྭ CHHWA, ཇྭ JWA, ཉྭ NHWA, ཉྫྭ NHJWA, ཏྭ TWA, ཐྭ THWA, དྭ DWA, དྷྭ DHWA, ནྭ NWA, ཌྭ DDWA, པྭ PWA, ཕྭ PHWA, བྭ BWA, བྷྭ BHWA, ཙྭ BBWA, ཡྭ YWA, ཧྭ RWA, ལྭ LWA, ཤྭ SSWA, སྭ SWA, ཧྭ HWA

5. Dependent vowel signs. The dependent vowels are as follows (shown with ཀ KA):

ཀ ka = ཀ ka
ཀཱ kā = ཀ ka + ◌ཱ -ā
ཀི ki = ཀ ka + ◌ི i
ཀཱི kī = ཀ ka + ◌ཱི -ī
ཀེ kei = ཀ ka + ◌ེ -ei
ཀུ ku = ཀ ka + ◌ུ -u
ཀཱུ kū = ཀ ka + ◌ུ -u + ◌ཱ -ā
ཀོ ko᷈ = ཀ ka + ◌ོ -o᷈
ཀཱོ ko᷈ = ཀ ka + ◌ོ -o᷈ + ◌ཱ -ā
ཀེོ ke = ཀ ka + ◌ེ -o + ◌ོ -o᷈
ཀཱེོ kē = ཀ ka + ◌ེ -o + ◌ོ -o᷈ + ◌ཱ -ā
ཀོ ko = ཀ ka + ◌ོ -o
ཀཱོ kō = ཀ ka + ◌ོ -o + ◌ཱ -ā
ཀཻ kai = ཀ ka + ◌ཻ -ai
ཀཽ kau = ཀ ka + ◌ེ -o + ◌ཽ -au
ཀ̣ ku̍ = ཀ ka + ◌̣ -u̍
ཀཱ̣ kū̍ = ཀ ka + ◌̣ -u̍ + ◌ཱ -ā

The relative order of the vowel signs as given here is the expected typing and storage order. In particular, U+AA29 VOWEL SIGN AA completes a vowel string.

6. Final consonants. Final consonants are indicated in three ways: by explicit final consonants with a long stroke (ཀ K, ང NG, ཉ NG, ཅ CH, ཏ T, ན N, པ P, ཡ Y, ར R, ལ L, ཤ SS), by combining marks (◌ྃ NG, ◌ཾ M, ◌ཿ H), and by ཧ, which functions as a final consonant without modification; compare ཅཅ *chach* and ཅཧ *chaw* with ཧཅ *vach* and ཧཧ *vaw*). Final consonants may occur word-internally: compare བྱཏརཀ *bitarak* with བྱཏརཀ *bitrak*—and indeed with བྱཏྲུ *bitruh*, which shows that there is a distinction between ཏར *tra* and ཏྲ *tra*. (The distinction is one of syllable

boundary, that is, *bit-rak* versus *bi-truh*.) The final ꨌ G is an innovation. The Cham academic community added this to eliminate some of the underdifferentiation that Cham writing has. Historically, final [k] and many final glottal stops were written with ꨋ FINAL K. Some Cham educators saw that ꨌ FINAL G did not occur in word-final position. So in their primers they use this letter to represent the final [k], leaving the ꨋ to represent many of the final glottal stops. See Figure 5.

7. Digits and punctuation. Digits have distinctive forms, though of course European digits are known because of Vietnamese usage. Three levels of break are recognized in Cham with ꩜ DANDA, ꩝ DOUBLE DANDA, and ꩞ TRIPLE DANDA, with progressive values of finality. The ꩉ SPIRAL often begins a section of text; the usage ꩜ꩉ at the beginning of a text has also been observed, as has the usage ꩝? to indicate a question. Nowadays, European marks like HYPHEN and COLON are also found. See Figures 6, 7, and 8.

8. Collating order. A syllabic ordering applies to Cham. The order given in Bùi Khánh Thế's 1995 Cham-Vietnamese dictionary is quite ideosyncratic and may not be suitable for the Common Tailorable Template. An order which treats initial letters (in binary order), then medial consonants as described in section 4, then vowel signs as listed in section 5, and then finals (in binary order), should give useful and predictable results. All sorting elements are treated with primary weight.

9. Character names. The character names used follow those given in Bùi Khánh Thế's 1995. Some transliterations may, of course, not reflect the pronunciation of a letter. AA27 ꨧ SA is pronounced the same as AA14 ꨔ THA, namely [tʰ], but this is a reading rule, and keeping the Brahmic-style transliteration names of Bùi Khánh Thế seems the most reasonable choice.

10. Linebreaking. Opportunities for line-break occur after any full orthographic syllable. Cham punctuation marks can be expected to have behaviour similar to that of Devanagari DANDA and DOUBLE DANDA.

11. Unicode Character Properties. Character properties are proposed here. With regard to the combining classes of the vowel signs, normalization should weight the combining marks in five groups:

M ◌ꩻ -YA, ◌ꩵ -RA, ◌ꩺ -LA, ◌ꩿ -WA
< P ◌ꩽ -O, ◌ꩾ -AI
< V ◌ꩰ -I, ◌ꩱ -II, ◌ꩲ -EI, ◌ꩳ -U, ◌ꩴ -UE, ◌ꩵ -AU, ◌ꩶ -UE
< L ◌꩷ AA
< F ◌꩸ -NG, ◌꩹ -M, ◌ꩺ -H

```
AA00;CHAM LETTER A;Lo;0;L;;;;;N;;;;;
AA01;CHAM LETTER I;Lo;0;L;;;;;N;;;;;
AA02;CHAM LETTER U;Lo;0;L;;;;;N;;;;;
AA03;CHAM LETTER E;Lo;0;L;;;;;N;;;;;
AA04;CHAM LETTER AI;Lo;0;L;;;;;N;;;;;
AA05;CHAM LETTER O;Lo;0;L;;;;;N;;;;;
AA06;CHAM LETTER KA;Lo;0;L;;;;;N;;;;;
AA07;CHAM LETTER KHA;Lo;0;L;;;;;N;;;;;
AA08;CHAM LETTER GA;Lo;0;L;;;;;N;;;;;
AA09;CHAM LETTER GHA;Lo;0;L;;;;;N;;;;;
AA0A;CHAM LETTER NGUE;Lo;0;L;;;;;N;;;;;
AA0B;CHAM LETTER NGA;Lo;0;L;;;;;N;;;;;
AA0C;CHAM LETTER CHA;Lo;0;L;;;;;N;;;;;
AA0D;CHAM LETTER CHHA;Lo;0;L;;;;;N;;;;;
AA0E;CHAM LETTER JA;Lo;0;L;;;;;N;;;;;
AA0F;CHAM LETTER JHA;Lo;0;L;;;;;N;;;;;
AA10;CHAM LETTER NHUE;Lo;0;L;;;;;N;;;;;
AA11;CHAM LETTER NHA;Lo;0;L;;;;;N;;;;;
AA12;CHAM LETTER NHJA;Lo;0;L;;;;;N;;;;;
AA13;CHAM LETTER TA;Lo;0;L;;;;;N;;;;;
AA14;CHAM LETTER THA;Lo;0;L;;;;;N;;;;;
AA15;CHAM LETTER DA;Lo;0;L;;;;;N;;;;;
AA16;CHAM LETTER DHA;Lo;0;L;;;;;N;;;;;
AA17;CHAM LETTER NUE;Lo;0;L;;;;;N;;;;;
```

```
AA18;CHAM LETTER NA;Lo;0;L;;;;;N;;;;;
AA19;CHAM LETTER DDA;Lo;0;L;;;;;N;;;;;
AA1A;CHAM LETTER PA;Lo;0;L;;;;;N;;;;;
AA1B;CHAM LETTER PPA;Lo;0;L;;;;;N;;;;;
AA1C;CHAM LETTER PHA;Lo;0;L;;;;;N;;;;;
AA1D;CHAM LETTER BA;Lo;0;L;;;;;N;;;;;
AA1E;CHAM LETTER BHA;Lo;0;L;;;;;N;;;;;
AA1F;CHAM LETTER MUE;Lo;0;L;;;;;N;;;;;
AA20;CHAM LETTER MA;Lo;0;L;;;;;N;;;;;
AA21;CHAM LETTER BBA;Lo;0;L;;;;;N;;;;;
AA22;CHAM LETTER YA;Lo;0;L;;;;;N;;;;;
AA23;CHAM LETTER RA;Lo;0;L;;;;;N;;;;;
AA24;CHAM LETTER LA;Lo;0;L;;;;;N;;;;;
AA25;CHAM LETTER VA;Lo;0;L;;;;;N;;;;;
AA26;CHAM LETTER SSA;Lo;0;L;;;;;N;;;;;
AA27;CHAM LETTER SA;Lo;0;L;;;;;N;;;;;
AA28;CHAM LETTER HA;Lo;0;L;;;;;N;;;;;
AA29;CHAM VOWEL SIGN AA;Mn;0;NSM;;;;;N;;;;;
AA2A;CHAM VOWEL SIGN I;Mn;0;NSM;;;;;N;;;;;
AA2B;CHAM VOWEL SIGN II;Mn;0;NSM;;;;;N;;;;;
AA2C;CHAM VOWEL SIGN EI;Mn;0;NSM;;;;;N;;;;;
AA2D;CHAM VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
AA2E;CHAM VOWEL SIGN OE;Mn;0;NSM;;;;;N;;;;;
AA2F;CHAM VOWEL SIGN O;Mc;0;L;;;;;N;;;;;
AA30;CHAM VOWEL SIGN AI;Mc;0;L;;;;;N;;;;;
AA31;CHAM VOWEL SIGN AU;Mn;0;NSM;;;;;N;;;;;
AA32;CHAM VOWEL SIGN UE;Mn;0;NSM;;;;;N;;;;;
AA33;CHAM CONSONANT SIGN YA;Mc;0;L;;;;;N;;;;;
AA34;CHAM CONSONANT SIGN RA;Mc;0;L;;;;;N;;;;;
AA35;CHAM CONSONANT SIGN LA;Mn;0;NSM;;;;;N;;;;;
AA36;CHAM CONSONANT SIGN WA;Mn;0;NSM;;;;;N;;;;;
AA40;CHAM LETTER FINAL K;Lo;0;L;;;;;N;;;;;
AA41;CHAM LETTER FINAL G;Lo;0;L;;;;;N;;;;;
AA42;CHAM LETTER FINAL NG;Lo;0;L;;;;;N;;;;;
AA43;CHAM CONSONANT SIGN FINAL NG;Mn;0;NSM;;;;;N;;;;;
AA44;CHAM LETTER FINAL CH;Lo;0;L;;;;;N;;;;;
AA45;CHAM LETTER FINAL T;Lo;0;L;;;;;N;;;;;
AA46;CHAM LETTER FINAL N;Lo;0;L;;;;;N;;;;;
AA47;CHAM LETTER FINAL P;Lo;0;L;;;;;N;;;;;
AA48;CHAM LETTER FINAL Y;Lo;0;L;;;;;N;;;;;
AA49;CHAM LETTER FINAL R;Lo;0;L;;;;;N;;;;;
AA4A;CHAM LETTER FINAL L;Lo;0;L;;;;;N;;;;;
AA4B;CHAM LETTER FINAL SS;Lo;0;L;;;;;N;;;;;
AA4C;CHAM CONSONANT SIGN FINAL M;Mn;0;NSM;;;;;N;;;;;
AA4D;CHAM CONSONANT SIGN FINAL H;Mc;0;L;;;;;N;;;;;
AA50;CHAM DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
AA51;CHAM DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
AA52;CHAM DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
AA53;CHAM DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
AA54;CHAM DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
AA55;CHAM DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
AA56;CHAM DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
AA57;CHAM DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
AA58;CHAM DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
AA59;CHAM DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
AA5C;CHAM PUNCTUATION SPIRAL;Po;0;L;;;;;N;;;;;
AA5D;CHAM PUNCTUATION DANDA;Po;0;L;;;;;N;;;;;
AA5E;CHAM PUNCTUATION DOUBLE DANDA;Po;0;L;;;;;N;;;;;
AA5F;CHAM PUNCTUATION TRIPLE DANDA;Po;0;L;;;;;N;;;;;
```

12. Bibliography.

Aymonier, Étienne and Antoine Cabaton. 1906. *Dictionnaire Čam-Français*. Paris.

Bùi Khánh Thế. 1995. *Từ diên Chăm-Việt:* ꨌꩌꨨꩍ ꨆꨯ꩓ꨚꨛꨩ ꨌꨯ. [Hồ Chí Minh]: Nhà xuất ban Khoa Học Xã Hội.

Inrasara. 2003. *Tự học tiếng Chăm*. Nhà xuất ban Văn hóa Dân Tộc.

Kōno Rokurō, Chino Eiichi, & Nishida Tatsuo. 2001. *The Sanseido Encyclopaedia of Linguistics. Volume 7: Scripts and Writing Systems of the World* [*Gengogaku dai ziten* (*bekkan*) *sekai mozi ziten*]. Tokyo: Sanseido Press. ISBN 4-385-15177-6

[s.n.] 1970. ꨆꨛꨨꨯꨚꨆꨛꨩꨚꨯ (Adei bac akhar Cham).

[s.n.] 1971. ꨐꨯꨩ ꨊꨛꨛ ꨆꨛꨩꨚꨯ ꨆꨯ (Thrơm bbak akhar Cham).

[s.n.] 1990. ꨆꨛꨚꨯ ꨙꨛꨚꨚꨆ ꨆꨯ꞉ ꨀꨛꨛ ꩌ (Akar panoc Cham: tag 2).

Acknowledgements
This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of the Cham encoding.

Figures

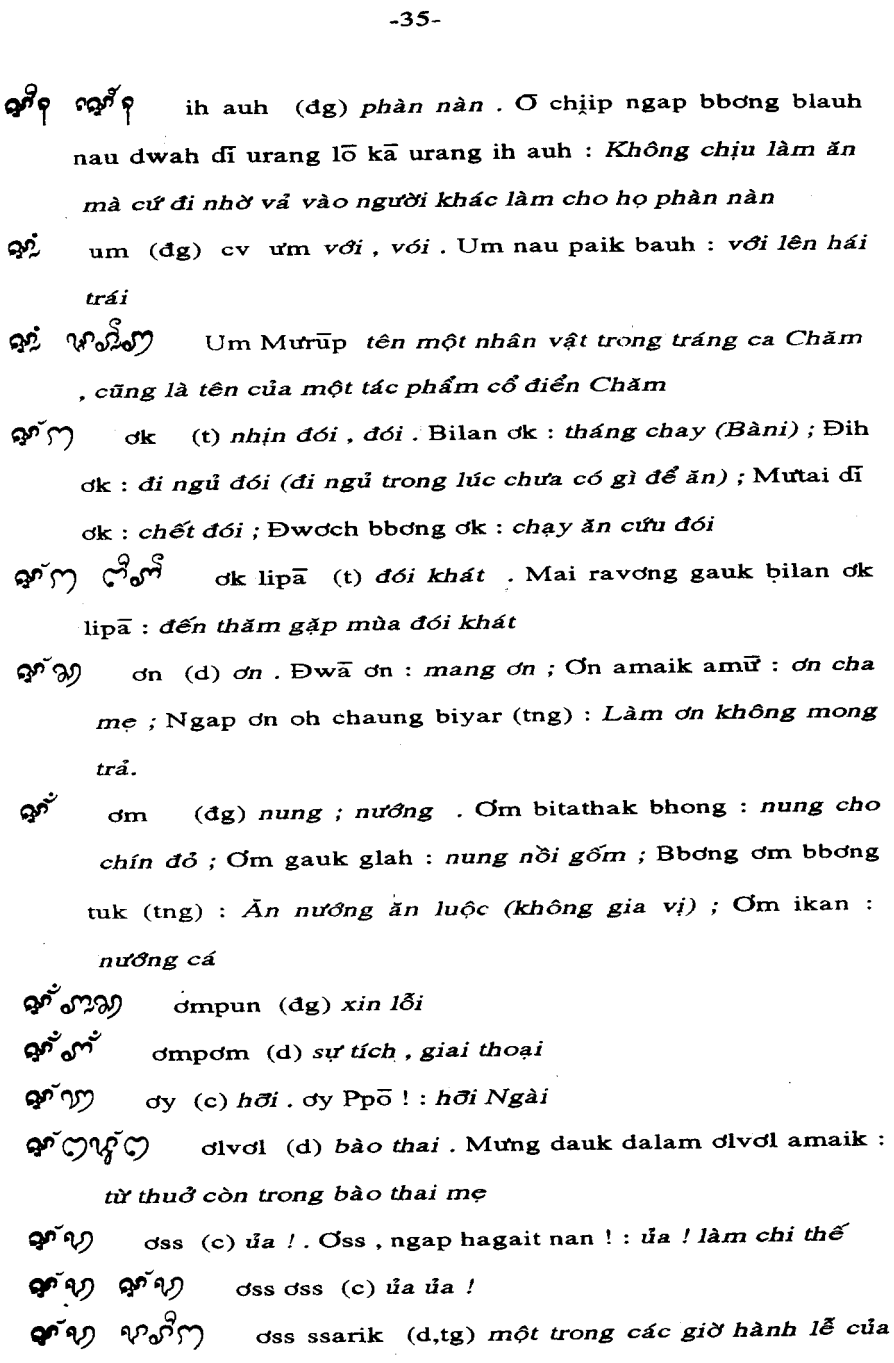
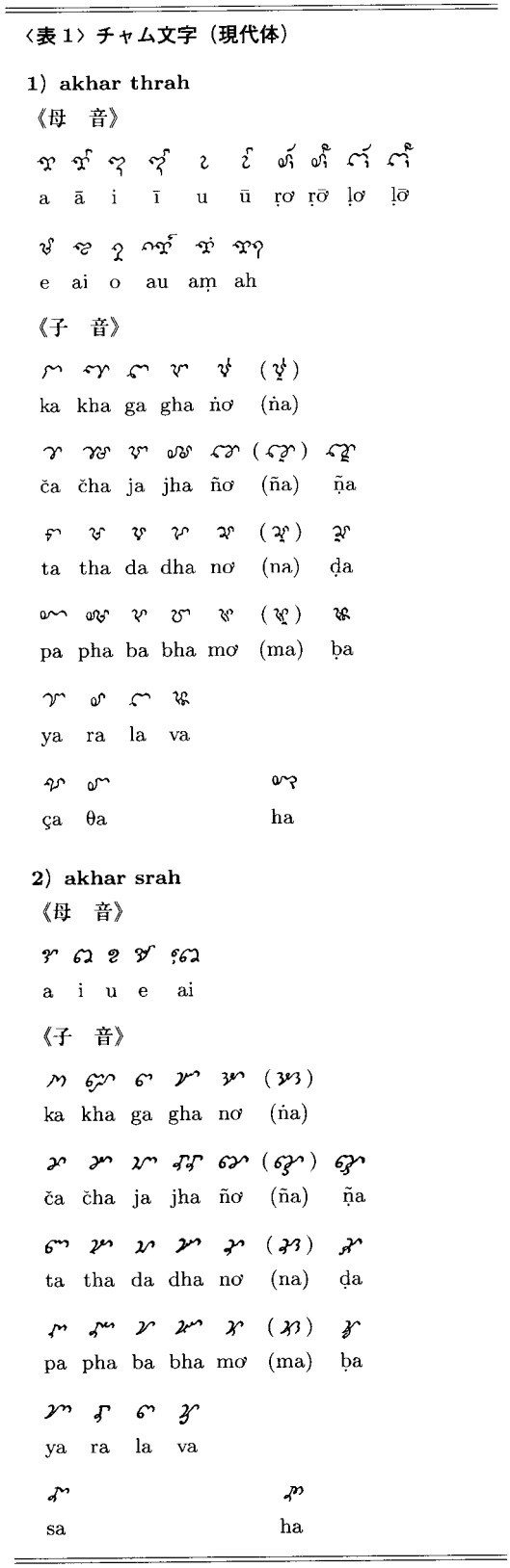


Figure 1. Sample from Bùi Khánh Thế's 1995 dictionary.



〈表 1〉チャム文字（現代体）

1) akhar thrah

《母　音》

a ā i ī u ū ṛơ ṝơ ḷơ ḹơ

e ai o au aṃ ah

《子　音》

ka kha ga gha ṅơ (ṅa)

ča čha ja jha ñơ (ña) ṇa

ta tha da dha nơ (na) ḍa

pa pha ba bha mơ (ma) ḅa

ya ra la va

ça θa ha

2) akhar srah

《母　音》

a i u e ai

《子　音》

ka kha ga gha nơ (ṅa)

ča čha ja jha ñơ (ña) ṇa

ta tha da dha nơ (na) ḍa

pa pha ba bha mơ (ma) ḅa

ya ra la va

sa ha

Figure 2. Sample from The Sanseido Encyclopaedia, showing different font styles for Cham.

Figure 3. Sample from a Cham reader showng the basic alphabet.

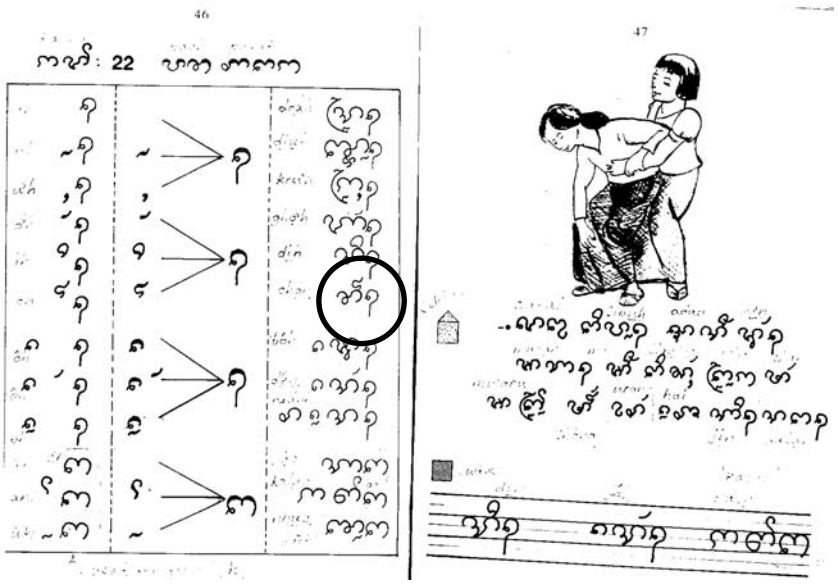


Figure 4. Sample from a 1990 Cham reader showing some vowel combinations with final *-h*, including ◌ꨯꩍ in ꨌꨯꩍ *coh* (which would contrast with ꨌꨌꨯꩍ **cauh*). It also shows ꨛ final *-g*.

Figure 5. Sample from a 1990 Cham reader showng some examples of ꨩ final *-g*. On the top right is ꨀꨔꨗꨀꨩ *uggamưg*.

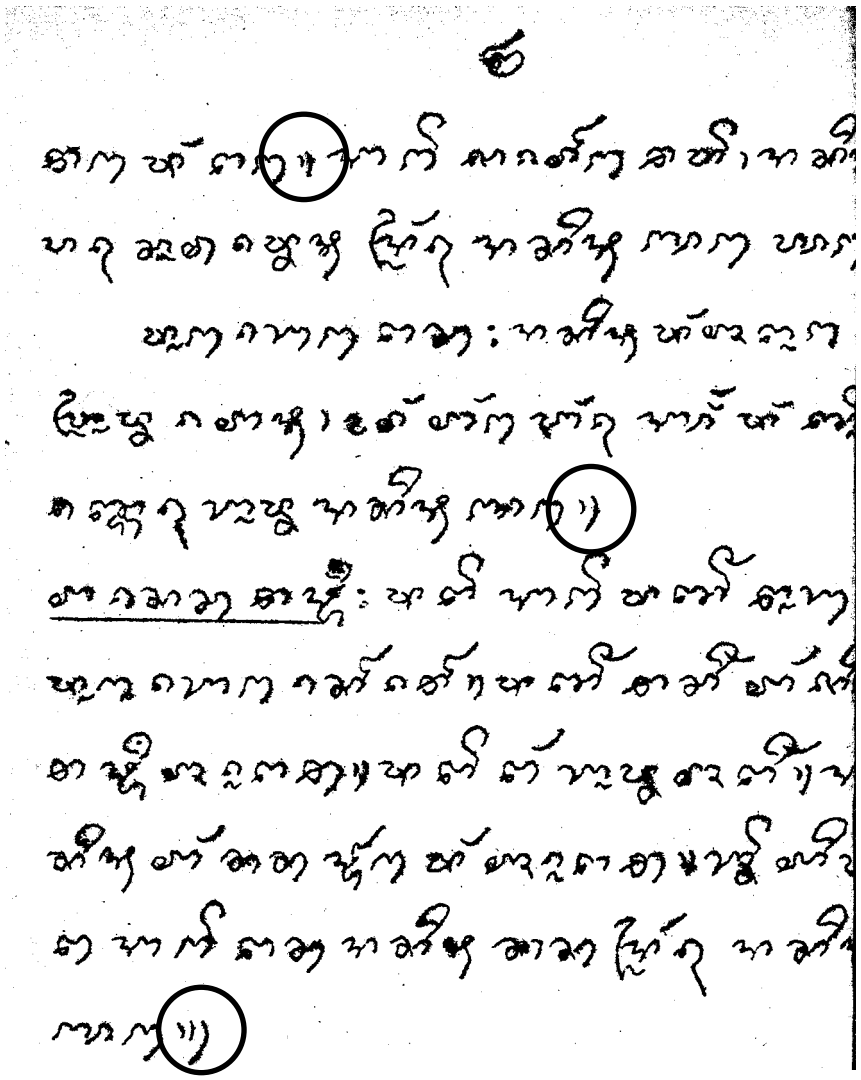


Figure 6. Sample from a Cham reader showing DOUBLE DANDA and TRIPLE DANDA.

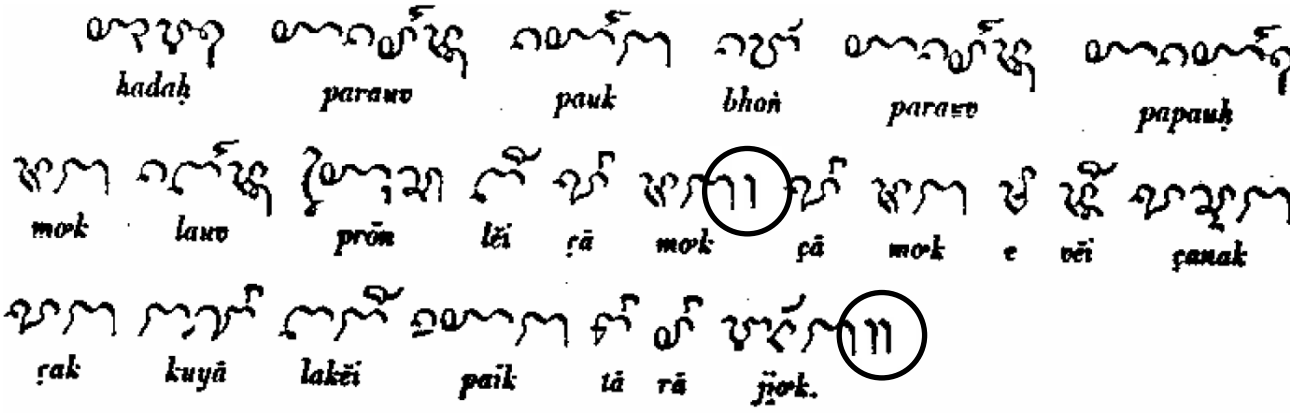


Figure 7. Sample from Aymonier and Cabaton 1906, showing DANDA and DOUBLE DANDA.

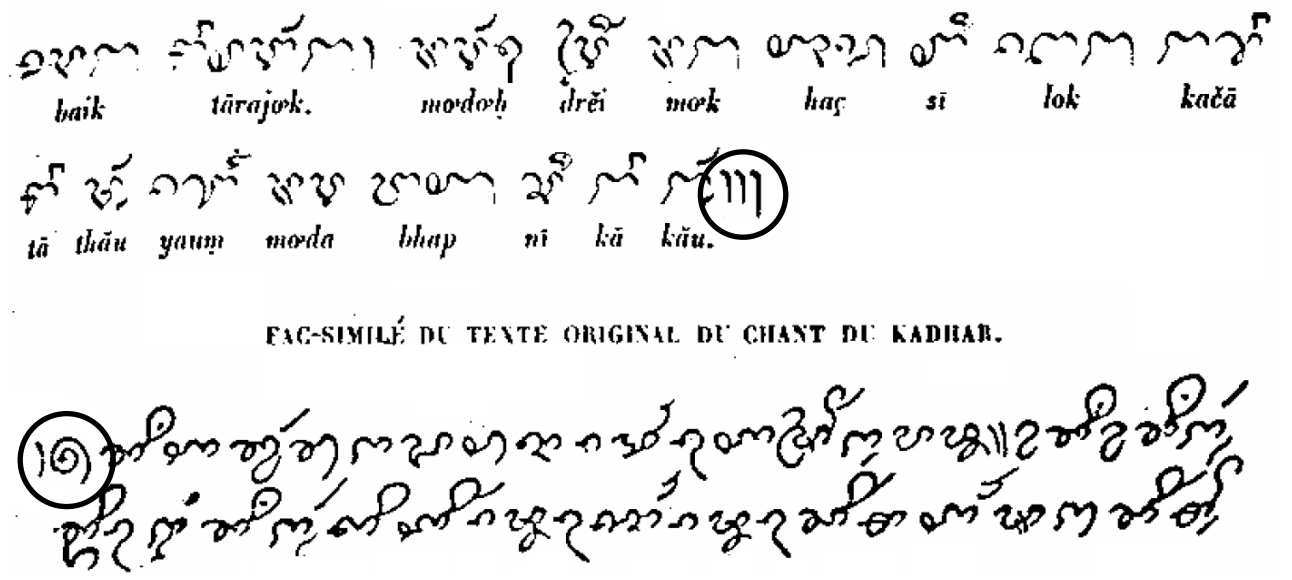


Figure 8. Sample from Aymonier and Cabaton 1906, showing TRIPLE DANDA, and DANDA, and SPIRAL.

TABLE xx - Row AA: CHAM

	AA0	AA1	AA2	AA3	AA4	AA5
0						
1						
2						
3						
4						
5						
6						
7						
8						
9						
A						
B						
C						
D						
E						
F						

G = 00
P = 00

TABLE XXX - Row AA: CHAM

hex	Name
00	CHAM LETTER A
01	CHAM LETTER I
02	CHAM LETTER U
03	CHAM LETTER E
04	CHAM LETTER AI
05	CHAM LETTER O
06	CHAM LETTER KA
07	CHAM LETTER KHA
08	CHAM LETTER GA
09	CHAM LETTER GHA
0A	CHAM LETTER NGUE
0B	CHAM LETTER NGA
0C	CHAM LETTER CHA
0D	CHAM LETTER CHHA
0E	CHAM LETTER JA
0F	CHAM LETTER JHA
10	CHAM LETTER NHUE
11	CHAM LETTER NHA
12	CHAM LETTER NHJA
13	CHAM LETTER TA
14	CHAM LETTER THA
15	CHAM LETTER DA
16	CHAM LETTER DHA
17	CHAM LETTER NUE
18	CHAM LETTER NA
19	CHAM LETTER DDA
1A	CHAM LETTER PA
1B	CHAM LETTER PPA
1C	CHAM LETTER PHA
1D	CHAM LETTER BA
1E	CHAM LETTER BHA
1F	CHAM LETTER MUE
20	CHAM LETTER MA
21	CHAM LETTER BBA
22	CHAM LETTER YA
23	CHAM LETTER RA
24	CHAM LETTER LA
25	CHAM LETTER VA
26	CHAM LETTER SSA
27	CHAM LETTER SA
28	CHAM LETTER HA
29	CHAM VOWEL SIGN AA
2A	CHAM VOWEL SIGN I
2B	CHAM VOWEL SIGN II
2C	CHAM VOWEL SIGN EI
2D	CHAM VOWEL SIGN U
2E	CHAM VOWEL SIGN OE
2F	CHAM VOWEL SIGN O
30	CHAM VOWEL SIGN AI
31	CHAM VOWEL SIGN AU
32	CHAM VOWEL SIGN UE
33	CHAM CONSONANT SIGN YA
34	CHAM CONSONANT SIGN RA
35	CHAM CONSONANT SIGN LA
36	CHAM CONSONANT SIGN WA
37	(This position shall not be used)
38	(This position shall not be used)
39	(This position shall not be used)
3A	(This position shall not be used)
3B	(This position shall not be used)
3C	(This position shall not be used)
3D	(This position shall not be used)
3E	(This position shall not be used)
3F	(This position shall not be used)
40	CHAM LETTER FINAL K
41	CHAM LETTER FINAL G
42	CHAM LETTER FINAL NG
43	CHAM CONSONANT SIGN FINAL NG
44	CHAM LETTER FINAL CH
45	CHAM LETTER FINAL T
46	CHAM LETTER FINAL N
47	CHAM LETTER FINAL P
48	CHAM LETTER FINAL Y
49	CHAM LETTER FINAL R
4A	CHAM LETTER FINAL L
4B	CHAM LETTER FINAL SS
4C	CHAM CONSONANT SIGN FINAL M
4D	CHAM CONSONANT SIGN FINAL H
4E	(This position shall not be used)
4F	(This position shall not be used)
50	CHAM DIGIT ZERO
51	CHAM DIGIT ONE
52	CHAM DIGIT TWO
53	CHAM DIGIT THREE
54	CHAM DIGIT FOUR
55	CHAM DIGIT FIVE
56	CHAM DIGIT SIX
57	CHAM DIGIT SEVEN
58	CHAM DIGIT EIGHT

hex	Name
59	CHAM DIGIT NINE
5A	(This position shall not be used)
5B	(This position shall not be used)
5C	CHAM PUNCTUATION SPIRAL
5D	CHAM PUNCTUATION DANDA
5E	CHAM PUNCTUATION DOUBLE DANDA
5F	CHAM PUNCTUATION TRIPLE DANDA

A. Administrative

1. Title
Proposal for encoding the Cham script in the BMP of the UCS.
2. Requester's name
UC Berkeley Script Encoding Initiative (Universal Scripts Project)
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2006-08-06
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.
1b. Proposed name of script
Cham.
1c. The proposal is for addition of character(s) to an existing block
No.
1d. Name of the existing block
2. Number of characters in proposal
83.
3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)
Category A.
4a. Proposed Level of Implementation (1, 2 or 3)
Level 2
4b. Is a rationale provided for the choice?
Yes.
4c. If YES, reference
Cham requires Level 2 implementation as other Brahmic scripts do.
5a. Is a repertoire including character names provided?
Yes.
5b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
5c. Are the character shapes attached in a legible form suitable for review?
Yes.
6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson.
6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes.
7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
8. Special encoding issues: Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?
Yes. See above.
9. Additional Information: Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.
Yes. See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes. See N1559, N1578, N1599, N1960.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.

2b. If YES, with whom?
Bùi Khánh Thế reviewed earlier versions of the proposal; Jim Brase and Dave Blood of SIL International have reviewed the proposal and made many useful suggestions.
2c. If YES, available relevant documents
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
Yes. See above.
4a. The context of use for the proposed characters (type of use; common or rare)
Common. Used to write the Cham language.
4b. Reference
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
Vietnam.
6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?
Yes. Positions AA00-AA5F are proposed.
6b. If YES, is a rationale provided?
Yes.
6c. If YES, reference
Contemporary use and accordance with the Roadmap.
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No.
8b. If YES, is a rationale for its inclusion provided?
8c. If YES, reference
9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
No.
10b. If YES, is a rationale for its inclusion provided?
10c. If YES, reference
11a. Does the proposal include use of combining characters and/or use of composite sequences?
Yes.
11b. If YES, is a rationale for such use provided?
Yes.
11c. If YES, reference
Brahmic vowel signs.
11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
No.
11e. If YES, reference
12a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
12b. If YES, describe in detail (include attachment if necessary)
13a. Does the proposal contain any Ideographic compatibility character(s)?
No.
13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?